LEVEL II

Research Report 376

AN MDI PROCEDURE FOR VULNERABILITY SEGMENTATION TESTS[1].

by

A./ Charnes
W.W./ Cooper
D.B./ Learner*
F.Y./ Phillips*

CENTER FOR CYBERNETIC STUDIES

The University of Texas
Austin, Texas 78712

DTIC
S ELECTE
DEC 1 7 1980
F
D

80 12 15 183

CCS-RR-376-REV

Research Report 376

AN MDI PROCEDURE FOR VULNERABILITY SEGMENTATION TESTS[1]

by

A. Charnes
W.W. Cooper
D.B. Learner*
F.Y. Phillips*

September 1980

[1]This paper represents a revision of an earlier version which appears as [4] in the bibliography.

CENTER FOR CYBERNETIC STUDIES

A. Charnes, Director
Business-Economics Building, 203E
The University of Texas at Austin
Austin, TX 78712
(512) 471-1821

ABSTRACT

This note describes a minimum discrimination information (MDI) method for testing the validity of the composite hypothesis involved in a "vulnerability segmentation," i.e., a set of groupings of products resulting from the vulnerability ratio algorithm of Charnes, Cooper, Learner, and Phillips. See [3]. The test is based on the large sample chi-square and multi-normal distribution properties of the functions involved and of the minimum discrimination information (MDI) statistic. An algorithm for effecting these tests is supplied which admits of straightforward implementation and computation via the well known and generally available SUMT program of Fiacco and McCormick.

KEY WORDS

Market Segmentation

Vulnerability Segments

Switching Constants

Statistical Tests

Minimum Discrimination Information

INTRODUCTION

This paper will detail a simple statistical test for the validity of product "vulnerability segmentations" as developed by Charnes, Cooper, Learner and Phillips [3]. The test involves (i) representing the hypothesized segmentation as a set of linear equations, and (ii) using the minimum discrimination information (MDI) statistic to determine whether an observed brand switching matrix is consistent with the hypothesized segmentation associated with these equations. Prior to detailing this test, we shall recapitulate the conclusions of the previous paper [3], and provide further comments on the meaning of the vulnerability ratios and vulnerability segments.

Vulnerability Ratios and Vulnerability Segments

The paper [3] made five main points. Three of these points may be summarized as follows:

(1) The "switching constants" of the Hendry literature (see [1], [6]) can be determined (for a product class) analytically and simultaneously, via a very simple one-pass algorithm which bypasses the iterated trial and error approaches on the basis of "expert judgment" and piecemeal calculation as suggested by others like Kalwani and Morrison for the Hendry system [8].[1]

(2) The switching constants are better interpreted as <u>vulnerability factors</u>. That is, the interpretation of these numbers as indices of "brand switching intensity" or "intensity of

---

[1]See also [1] and [6].

1

competition" is only partially justified." The vulnerability factor (or ratio) reflects the instability of the brand's share of market in the form of susceptibility to share losses from competitive action.

(3) It follows that the brands in a category can be segmented according to similar values of their observed vulnerability ratios. This segmentation can be tested for validity using information theoretic statistics. See [3] and [7].

Such segments need not involve groupings of brands that "compete most intensely with each other," as is sometimes assumed. They consist rather of equally vulnerable products or, from another point of view, products with equal share growth potential.

We shall elaborate further on this below, after we first record the other two points in the five to be summarized from [3] as follows:

(4) Knowing the pattern of segmentation is of limited managerial value until we also know what dimensions or attributes make one product preferable to another. This means that other models of brand choice behavior must be called in to round out the picture.

(5) There has been scholarly controversy over the rule of "switching proportional to brand share" [1] , [6]. Can this law be discerned empirically or must it be assumed or derived theoretically? Does it require an "entropic" theory a la Hendry, or the "zero-order process" assumption of Kalwani and Morrison? We show in [3] that the simplest possible assumption about multi-brand behavior (i.e., a multinomial model) leads to switching proportional to market share as a result of the large-sample covariance matrix of brand purchases. This is the same result that enables us to construct the statistical tests of the segmentation pattern that we will develop later in this paper.

The concept of brand vulnerability ratios can have relevant and material impact on marketing management analysis, and the following comments amplify this notion: Because the idea of brand switching is of central importance in marketing, many analysts assume that the non-repeat, off-diagonal purchases on a cell by cell basis in a switching table represent competitive

interaction among brands, specifically between pairs of brands. The assumption that switching between pairs of brands describes competition between those brands has been summarized in the notion of a "switching constant," a supposed measure of competition between brand pairs. Upon analysis as in the next section of this paper, however, this so-called "constant" is a ratio of non-repeat purchase to expected non-repeat purchase. In essence, it is a ratio of the number who go away to the number one expects would go away. This is a consumer out-flow measure, so that the value of the ratio is a measure of brand or product vulnerability to such outflows. Hence, our use of the name "vulnerability ratio."

The vulnerability ratio can be interpreted over its range from zero through 1.0 and beyond in ways which we shall subsequently justify in detail. In anticipation of this justification, we summarize as follows: A vulnerability ratio of zero means that the brand is invulnerable to the marketing efforts of other brands. That is, "invulnerability" means 100% repeat buying. In effect, trial is equal to repeat. Clearly, if one were so fortunate as to manage a brand of this variety, one would have to raise a question of how meaningful it is to continue marketing activity at current levels, since such activity is not likely to improve repeat purchase behavior of current consumers.[1]

A vulnerability ratio of 1.0 reflects random repeat purchase behavior. To the extent that consumers reflect such random behavior, an analyst should raise the question of marketing effectiveness since such random repeat purchases means that purposeful marketing activities lead merely to random consumer behavior.

---

[1]One might, however, justify continuing such activity to increase the frequency with which the brand is purchased.

To the extent that the vulnerability ratio exceeds 1.0 the brand exhibits attributes that tend to discourage repeat buying. Some combination of brand characteristics then discourages repeat purchases. It is not clear how frequently such greater than 1.0 circumstances may be encountered. To the extent that it prevails, however, it indicates that maximum vulnerability occurs when repeat purchase is zero.

Any brand must derive its growth from new buyers or other brands' non-repeaters. The vulnerability ratio identifies those brands with larger non-repeat buying. Thus the flow of non-repeat buyers can be effectively monitored through SANDDABS or like analyses of brand switching (and repeat buying) behavior.[1] It should then be clear that when a brand manager is aware of the vulnerability ratio of his brand and others, he must deal with the management question of what his brand's vulnerability ratio should be, and this involves questions of establishing management goals and developing plans to achieve those goals.

Brands may be clustered to form market segments on the basis of their vulnerability. Historically, analysts have ignored the question of whether one segmentation is significantly different from another.

This oversight is a major shortcoming of segmentation methodology. It is not difficult to find a hierarchical structure that adequately fits a set of data, but many other hierarchies might fit equally well. There is, therefore, need for a statistical test that discriminates between alternative structures. Note, however, that such tests must generally refer to composite rather than only simple hypotheses since the possible segmentations must be tested simultaneously and not in a separate fashion. This then enters into the choice of

---

[1]See Charnes, Cooper, and Learner [2].

the MDI (minimum discrimination information) statistic that we shall use.[1]
Our vulnerability ratio approach then provides a simple, direct, and unique
way, as we shall see, for determining whether one hypothesized segmentation
is significantly different from other possible segmentations. This capability
provides the analyst and the brand manager with added confidence that they
are dealing with a stable reality.

In summary, the vulnerability ratio allows us to order the brands in
a product category according to their vulnerability to competing brands and
to segment the brands on the basis of evidence in an easy and straightforward
manner by means of a one-pass algorithm provided via the SSI theorem (and
supporting lemmas) in [3]. Statistical testing by means of the MDI statistic
to accord with desired degrees of statistical confidence can then be effected
as we shall show in this paper or, alternatively, the sample size needed to
achieve these confidence levels can be ascertained by means of a simple
algorithm applied to the models for implementing these tests that we shall
also provide.

In this way, we can provide a phased or an integrated approach,
according to whether significance testing is wanted for guidance to marketing
goals directed to changing or taking advantage of possible vulnerabilities.
These are desirable capabilities for marketing managers.

INTERPRETATIONS OF THE VULNERABILITY RATIO

Our previous paper [3] described the new interpretations we have
given to the "switching constant"--e.g., in the literature on Hendry type
analyses--and our reasons for renaming it a _vulnerability ratio_. We concurred

---

[1]A detailed discussion of uses of the MDI statistic for testing purposes
with accompanying justifications and explanations may be found in [7]. Its uses
for unifying many other statistical models (and methods) now being or suggested
for use in marketing is described in [5].

with the Hendry idea that marketing insights can be gained by grouping to-
gether products having similar values of the constant, but concluded that
these _vulnerability segments_ must also be given somewhat different interpre-
tations. We develop these interpretations more explicitly in the section
that follows.

INTERPRETATIONS

The vulnerability ratio for brand i is computed by the formula

$$R_i = \frac{1 - \frac{(\text{repeat rate of i})}{(\text{share of i})}}{1 - (\text{share of i})}$$

See [3]. The smaller the ratio, the more stable, or invulnerable, the brand.
It may help if we also observe that R_i may be regarded as an extension of
the simpler ratio

$$\frac{\text{repeat share}}{\text{brand share}} ,$$

but R_i makes additional properties evident which can be put to good use in
marketing management. For example, R_i is favorable within the range of 0 to 1,
with the former (zero value) representing the more favorable limit of this
favorable range. Indeed, if $R_i=0$, then simple algebra shows that

Brand share = repeat rate.

That is, _all_ buyers of the brand are repeat buyers. The brand is invulnerable
to competitive luring away of its customers and it will keep any new customers
it attracts.

At the part of the range where $R_i=1$, we find

repeat rate = (brand share)^2.

This means that successive purchases of brand i are statistically independent
events, which is to say that repeat buying of i is a random phenomenon.

In general, then, a value of R_i strictly between 0 and 1 means that brand i has some favorable characteristics that turns "triers" into "users." If R_i is greater than 1, then repeat buying is even less than one would expect from random purchasing. Something is discouraging triers from becoming users. If brand i has an R_i greater than 1 and a large market share, we conclude that most of i's customers are new triers. For a given market share, R_i reaches a maximum when the repeat rate is zero.

This kind of information can be put to use in developing market strategies. For instance, it directs attention to the possibilities of attracting customers from brands with R_i values exceeding one's own. Conversely, it warns of possible inroads from brands with smaller R_i values. A mere finding of differences in the R_i values computed in accordance with [3], however, is not sufficient. Statistical (and other) types of validation may also be required and it is to the statistical aspects of this task that we now turn.

AN ALGORITHM FOR DEVELOPING VULNERABILITY SEGMENTS

The SSI theorem in [3] permits us to initiate our tests via hypothesized equalities computed from

$$(1) \qquad R_i = \frac{1 - K_{ii}/P_i}{1 - P_i}$$

where $K_{ij} \geq 0$ is an element of the observed switching matrix with $\sum_{i,j} K_{ij} = 1$ and P_i is the market share of brand i. These are the same R_i which were introduced in the SSI theorem of [3]. We have, however, replaced the p(i,i) of that theorem with the symbols K_{ii} and the related K_{ij} to underscore that these values are to be treated as constants in the analysis that we now introduce.

We will use the following canonical model to illustrate our procedure
for testing the structure of a market in the form of series of nested
hypotheses[1]

$$\text{Minimize} \quad I(p_{ij}:K_{ij}) \equiv \sum_{i,j} p_{ij} \ln \frac{p_{ij}}{K_{ij}}$$

subject to

$$(2) \qquad 1 = \sum_{i,j} p_{ij}$$

$$0 = \hat{R}_i - \hat{R}_j ,$$

Solving the above problem[2] yields values
where $p_{ij} \geq 0$ all i, j.

$$(3) \qquad \hat{R}_i^* = \frac{1 - p_{ii}^*/p_i}{1 - p_i}$$

$$\hat{R}_j^* = \frac{1 - p_{jj}^*/p_j}{1 - p_j}$$

with $\hat{R}_i^* = \hat{R}_j^*$ for all pairs which are hypothesized to be in the same
segment. The choices p_{ij}^* obtained from

$$(4) \qquad M \text{ in } I \ (p_{ij} :K_{ij}) = I \ (p_{ij}^* : K_{ij})$$

---

[1]As noted in [5], "Gokhale and Kullback [6] explain procedures for
testing nested hypotheses with MDI and Phillips [9] provides additional
examples. Nested hypotheses are effected by adding or removing constraints
in the canonical MDI problem; the associated information values and degrees
of freedom are additive and can be displayed in an 'Analysis of Information'
table. The sequential procedure is both convenient and meaningful. Further-
more, I(p:K) is a "general distance measure" and not merely a "test statistic,"
and so even when an hypothesis is rejected, the MDI test procedure will then
determine the best alternative hypothesis."

[2]The well know SUMT program of Fiacco and McCormick is available for this
purpose.

which also satisfy the constraints gives the MDI statistic which we shall symbolize by

$$(5) \quad I^* \equiv I\left(p_{ij}^* : K_{ij}\right).$$

This is the statistic that we shall use to test the hypothesized vulnerability segments in the following manner.

Ordering the R_i values from smallest to greatest as $R_{(1)}$, $R_{(2)}$, ..., $R_{(n)}$, we now suggest the following algorithm as a systematic way for using MDI tests of nested linear hypotheses to determine a vulnerability segmentation of the market. The proposed algorithm proceeds as follows: By MDI, test the hypothesis that $R_{(1)}$ and $R_{(2)}$ are equal at a specified level of significance. If the resulting value of the MDI statistic rejects the hypothesized equality then segment $R_{(1)}$ from $R_{(2)}$. Next begin with $R_{(2)}$ as a new smallest R_i to test for a further segment apart from $R_{(3)}$. If, instead, the test accepts the hypothesized equality of $R_{(1)}$ and $R_{(2)}$, add the additional condition to $R_{(1)} = R_{(2)}$ that $R_{(2)} = R_{(3)}$. Test with the MDI statistic. If the test accepts the hypothesized equality of $R_{(2)}$ and $R_{(3)}$ go on to new higher R_i by adding the new pertinent additional equality conditions.

To illustrate our suggested procedures for using these MDI values, we now refer to the following R_i values calculated from the Ehrenberg-Goodhardt data as explained in [3]:

i:	2	1	6	4	3	5
R_i:	.656	.660	.704	.719	.723	.750

For brevity and clarity, we shall present for illustration only the penultimate and ultimate MDI tests within the algorithm we have set forth above.

Given that R_2 and R_1 have been assigned to one segment separate from R_6 and R_4 in previous applications of the above algorithm, the penultimate test involves the possible addition of i=3 to the segment tentatively composed of i=6 and 4 from the values in the preceding table. For a confidence level of 0.95 and N=185 we calculate both this and the ultimate possible addition of i=5 to obtain the following I* and related values:

Segmentation Test	I*	2NI*	d.f.	$\chi^2_{0.95}$
3 added	.0208	7.696	3	7.815
3 and 5 added	.02122	7.844	4	9.488

Here we are assuming that 2NI* is distributed as χ^2, which we are entitled to do by virtue of the statistical considerations set forth in the next-to-last section of [3]. The $\chi^2_{0.95}$ values exhibited in this table then represent the maximum acceptance (i.e., failure to reject) levels under each of the 2 indicated hypotheses. We may also note that the above I* values were calculated from the following problems which were formed from (3) and which have so-called "nested" linear constraints in the p_{ij}:

$$\text{minimize} \quad \sum_{i,j} p_{ij} \ln \frac{p_{ij}}{K_{ij}}$$

subject to

$$1 = \sum_{i,j} p_{ij}$$

$$0 \leq p_{ij}, \qquad \text{all } i,j$$

$$(4) \qquad 0 = \hat{R}_2 - \hat{R}_1$$

$$0 = \hat{R}_4 - \hat{R}_6$$

$$0 = \hat{R}_3 - \hat{R}_4$$

and

$$\text{minimize} \quad \sum_{i,j} p_{ij} \ln \frac{p_{ij}}{K_{ij}}$$

subject to

$$1 = \sum_{i,j} p_{ij}$$

$$0 \leq p_{ij} \qquad \text{all } i,j$$

(5)
$$0 = \hat{R}_2 - \hat{R}_1$$

$$0 = \hat{R}_4 - \hat{R}_6$$

$$0 = \hat{R}_3 - \hat{R}_4$$

$$0 = \hat{R}_5 - \hat{R}_3$$

Since $2NI^* \leq \chi^2_{0.95}$ for both the addition of 3 and the addition of 3 and 5 to the second segment, we conclude that there are only 2 segments {2,1} and {6,4,3,5} for samples in the vicinity of N=185 and confidence level 0.95 rather than the 3 segments that were hypothesized in [3]. As is evident from the values for $2NI^*$ in the above table, these tests depend on N and, in fact, for $N \leq 40$ the Ehrenberg-Goodhardt hypothesis of no segmentation is valid. See [3]. Other possibilities at other sample sizes and significance levels are evidently also possible, but we do not propose to examine them here.[1]

CONCLUSION

The above characterizations are intended to supply justification for the use we have made of the MDI statistics for testing these composite hypotheses. There are also other reasons for using this statistic, however, in that we have elsewhere [5] shown how this information theoretic concept can be used to unify logit and probit analyses with still other approaches

---

[1]For further discussion see [4].

to a wide variety of marketing problems. The initial promise of the use

of this statistic for such unification in the course of exploring and

extending the original SANDABS analyses for brand switching with changing

market size as formulated in [2] has thus been carried forward yet another

step in the course of the above application for the composite hypothesis

testing associated with the segmentation analyses developed in this note.[2]

---

[2] See also [5].

BIBLIOGRAPHY

1. Butler, D. and B. Butler Jr., "Hendrodynamics: Fundamental Laws of
    Consumer Dynamics". Cambridge, Mass., The Hendry Corporation,
    Chapter 1, 1970, and Chapter 2, 1971.

2. Charnes, A., W. W. Cooper, and D. B. Learner, "Constrained Information
    Theoretic Characterizations in Consumer Purchase Behavior."
    J.Opl.Res.Soc., 29, 1978, pp. 832-842.

3. Charnes, A., W. W. Cooper, D. B. Learner and F. Y. Phillips, "A
    Theorem and Approach to Market Segmentation". Market Research
    Corporation of America PDA-RP#23, 1980.

4. Charnes, A., W. W. Cooper, D. B. Learner and F. Y. Phillips, "MDI
    Procedure for Vulnerability Segmentation", Market Research Cor-
    poration of American PDA-RP#120, July, 1980.

5. Charnes, A., W. W. Cooper, D. B. Learner and F. Y. Phillips, "The
    MDI Method as a Generalization of Logit, Probit and Hendry
    Models in Marketing". Market Research Corporation of America
    PDA-RP#70, 1980.

6. Ehrenberg, A.S.C., and G. J. Goodhardt, "The Hendry Brand Switching
    Coefficient". ADMAP, August 1974, pp. 232-238. Also available
    from London, ASKE Research Ltd.

7. Gokhale, D. V. and S. Kullback, The Information in Contingency Tables
    New York: Marcel Dekker, Inc., 1978.

8. Kalwani, M., and D. G. Morrison, "A Parsimonious Description of the
    Hendry System". Management Science, 23, No. 5, January 1977.

9. Phillips, F. "A Guide to MDI Statistics for Planning and Management
    Model Building." Austin: University of Texas at Austin, Insti-
    tute for Constructive Capitalism, 1980.

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author)	2a. REPORT SECURITY CLASSIFICATION
Center for Cybernetic Studies The University of Texas at Austin	Unclassified
	2b. GROUP

3. REPORT TITLE

An MDI Procedure for Vulnerability Segmentation Tests

4. DESCRIPTIVE NOTES *(Type of report and inclusive dates)*

5. AUTHOR(S) *(First name, middle initial, last name)*

A. Charnes, W.W. Cooper, D.B. Learner, F.Y. Phillips

6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
September 1980	14	9

8a. CONTRACT OR GRANT NO.	9a. ORIGINATOR'S REPORT NUMBER(S)
N00014-75-C-0569	
b. PROJECT NO.	CCS 376
NR047-021	
c.	9b. OTHER REPORT NO(S) *(Any other numbers that may be assigned this report)*
d.	

10. DISTRIBUTION STATEMENT

This document has been approved for public release and sale; its distribution is unlimited.

11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY
	Office of Naval Research (Code 434) Washington, DC

13. ABSTRACT

This note describes a minimum discrimination information (MDI) method for testing the validity of the composite hypothesis involved in a "vulnerability segmentation," i.e., a set of groupings of products resulting from the vulnerability ratio algorithm of Charnes, Cooper, Learner, and Phillips. See [3]. The test is based on the large sample chi-square and multi-normal distribution properties of the functions involved and of the minimum discrimination information (MDI) statistic. An algorithm for effecting these tests is supplied which admits of straightforward implementation and computation via the well known and generally available SUMT program of Fiacco and McCormick.

14 KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
Market Segmentation						
Vulnerability Segments						
Switching Constants						
Statistical Tests						
Minimum Discrimination Information						